Supplementary material for
The rupture extent of low frequency earthquakes near Parkfield, CA

Jessica C. Hawthorne (Department of Earth Sciences, University of Oxford, Oxford, UK)
Amanda M. Thomas (Department of Earth Sciences, University of Oregon, Oregon, USA)
Jean-Paul Ampuero (Université Côte d'Azur, IRD, CNRS, Observatoire de la Côte d'Azur, Géoazur, France; California Institute of Technology, Seismological Laboratory, Divisional of Geological and Planetary Sciences, Pasadena, CA, USA))

S1 Writing P_c as a Sum Over Stations

In equation (7), we compute the inter-station coherent power averaged over station pairs:

$$P_c = \frac{2}{N(N-1)} \sum_{k=1}^{N} \sum_{l=k+1}^{N} a_k a_l \operatorname{Re}\left(\hat{x}_k \hat{x}_l^*\right) \tag{S1}$$

We note, however, that this value can be rewritten as an average over stations to speed up the calculations. Following *Hawthorne and Ampuero* (2017) (section S2),

$$\begin{aligned} P_c &= \frac{1}{N(N-1)} \sum_{k=1}^{N} \sum_{l \neq k} a_k a_l \operatorname{Re}\left(\hat{x}_k \hat{x}_l^*\right) && \text{(S2)} \\ &= \frac{1}{N(N-1)} \left[\operatorname{Re}\left(\sum_{k=1}^{N} a_k \hat{x}_k \sum_{l=1}^{N} a_l \hat{x}_l^* \right) - \sum_{k=1}^{N} a_k^2 \operatorname{Re}\left(\hat{x}_k \hat{x}_k^*\right) \right] && \text{(S3)} \\ &= \frac{1}{N(N-1)} \left[\left| \sum_{k=1}^{N} a_k \hat{x}_k \right|^2 - \sum_{k=1}^{N} a_k^2 |\hat{x}_k|^2 \right]. && \text{(S4)} \end{aligned}$$

S2 LFE Stacks and Amplitudes

When we create stacked templates for each LFE family in section 4.2, we normalize the individual records by their maximum amplitude and weight by the station-averaged cross-correlation obtained by *Shelly* (2017). Then we isolate the stations and channels that have a signal to noise ratio of at least 3 in the 2 to 10 Hz band. These stacks provide the time dependence of the template's source time function and Green's function, but their amplitudes are not meaningful. Therefore we rescale the templates to match the amplitudes seen in individual LFE records.

For each LFE record, we bandpass filter to 2 to 8 Hz, compute the unnormalized cross-correlation of the LFE record with a 3-second-long portion of the template ($d_{jk} \cdot d_{tk}$) and then divide by the templates autocorrelation ($d_{tk} \cdot d_{tk}$). The set of amplitudes $(d_{jk} \cdot d_{tk})/(d_{jk} \cdot d_{tk})$ contains the factors required to scale the template seismogram to match the individual observations. As a first estimate of the appropriate template scaling factor, we take the median scaling factor on each channel.

However, these factors could be biased, as the LFE moment can vary from event to event, and not all LFEs are observed at all stations. We then use the relative amplitudes observed with the 20-odd channels to estimate the relative moments of each LFE. For each LFE, we divide the observed amplitudes by the channel-dependent scaling factor. We take the median among the channels as an estimate of the relative LFE moment.

We use these LFE moments to adjust the amplitude scaling for each station. If the moments are correct, the ratio of the observed amplitudes to the moments should be constant among the several thousand LFEs. So we take the median ratio as a new estimate of the appropriate scaling factor. In this calculation, we use only LFEs that have data available for at least 70% of the well-resolved channels. Next, we iterate, recomputing the relative moments and the scaling factors four or five times until the values converge.

Finally, we iterate the stacking with subsets of larger-amplitude LFEs to slightly improve the signal to noise ratio. Our initial stack uses all available LFE observations. In subsequent stacks, we use only records that have amplitudes between 0.2 and 6 times the median amplitude for that station. We recompute the stacks and scaling factors 4 times, at which point the values have converged.

The stacked templates obtained via this method are shown in panel a of Figures S1 to S7. Panels b show the amplitude spectra of 3-second-long portions of the LFE templates in comparison with the spectra of "noise" intervals, taken 4.1 to 1.1 s before the LFE S arrival. Note that this interval often includes the P arrival, which is effectively noise for our coherence analysis. Panels c show the signal to noise ratio. We require that the average signal to noise amplitude ratio in the 2 to 10 band be at least 3 (gray bars).

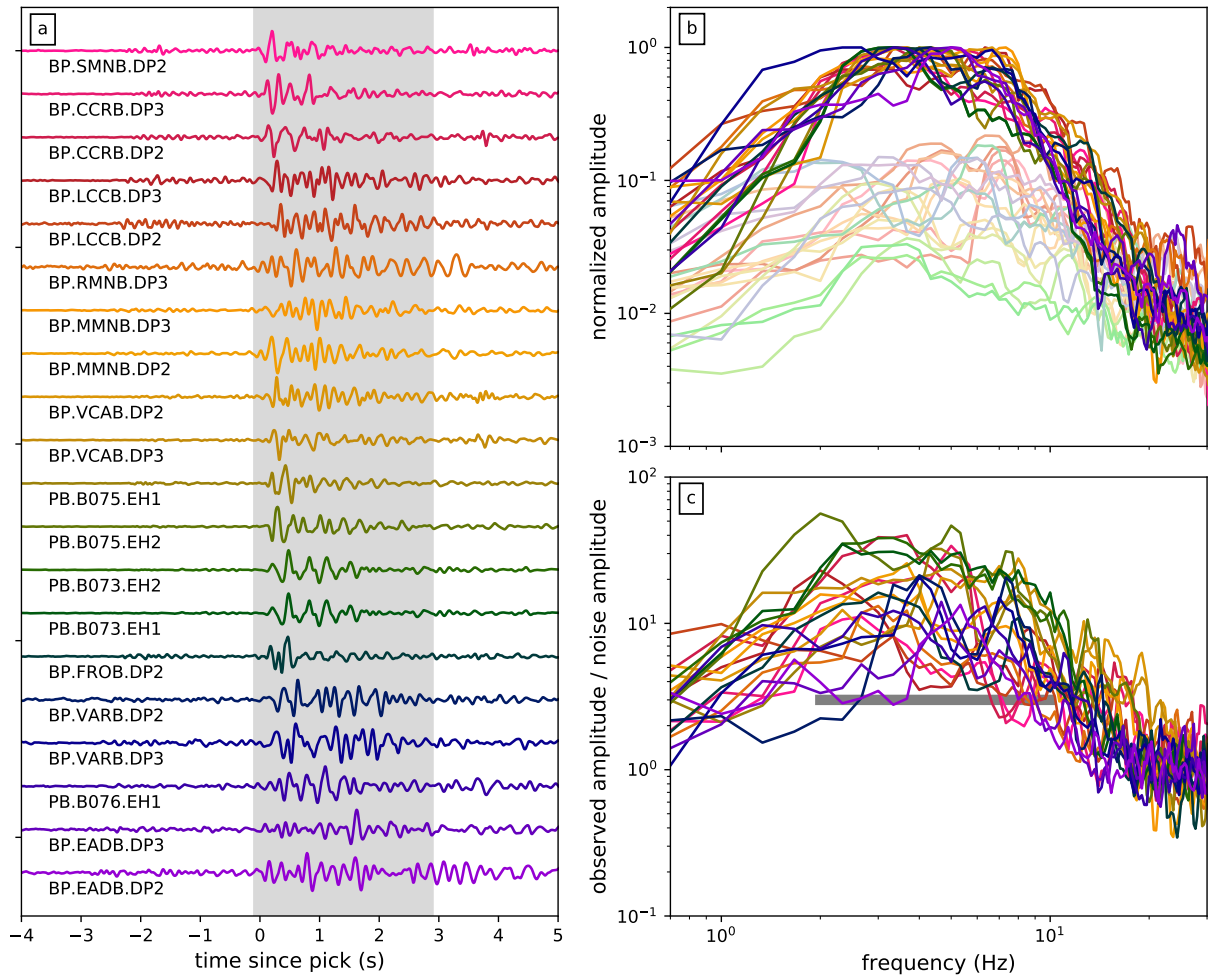

Figure S1: (a) Stacked templates for LFEs in Family 37140. Text indicates the network, station, and channel. The gray bar indicates the time intervals used for the signal to noise check and in the power calculations. (b) Dark lines: Normalized amplitude spectra of the LFE intervals. Pastel lines: Normalized amplitude spectra of noise, from 3-second intervals starting 4 s before the LFE intervals. (c) Ratio of the LFE amplitude to the noise amplitude. The horizontal gray bar indicates signal amplitude we require to use a given template, as averaged over the marked frequency range.

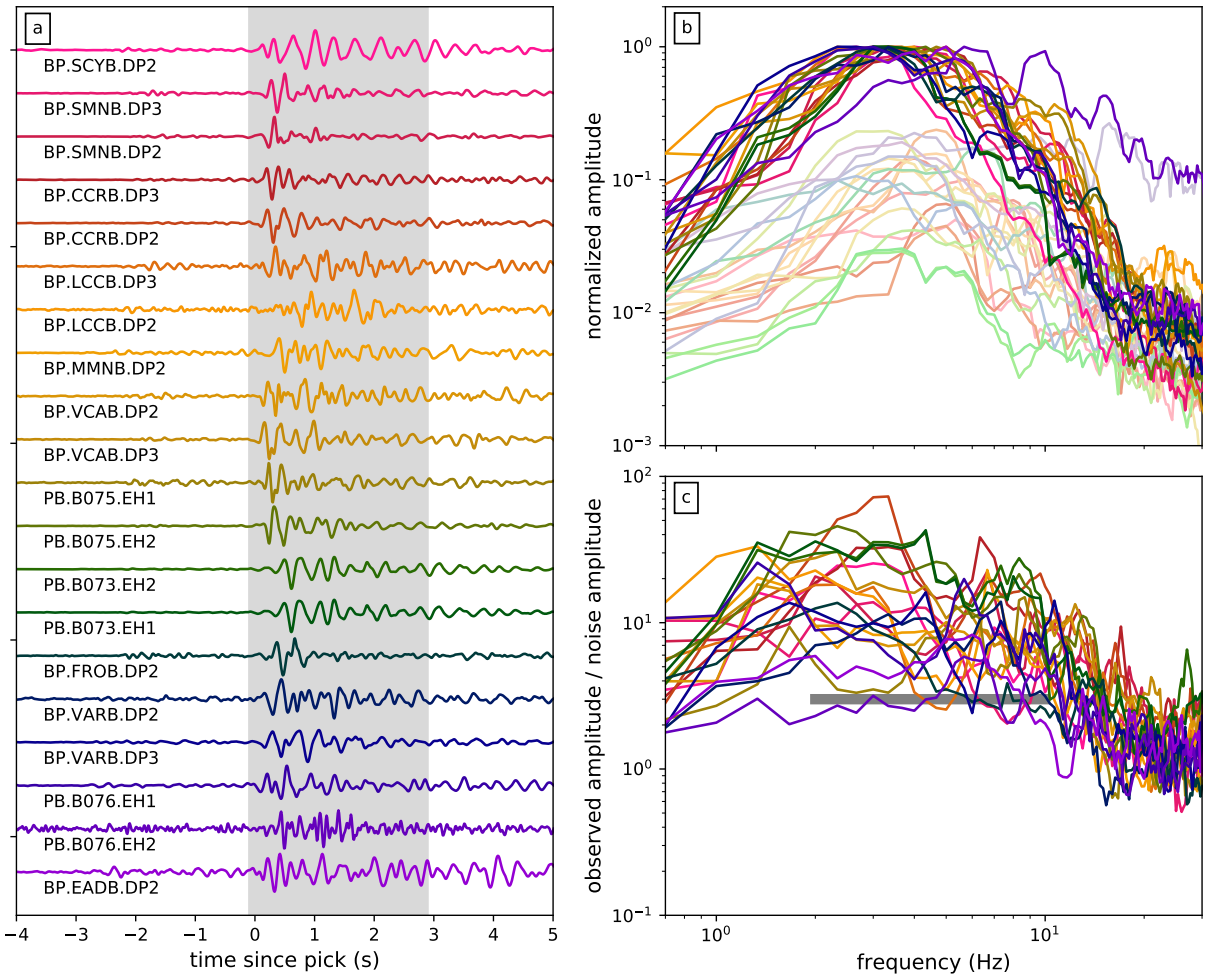

Figure S2: As in Figure S1, but for LFEs in family 37102.

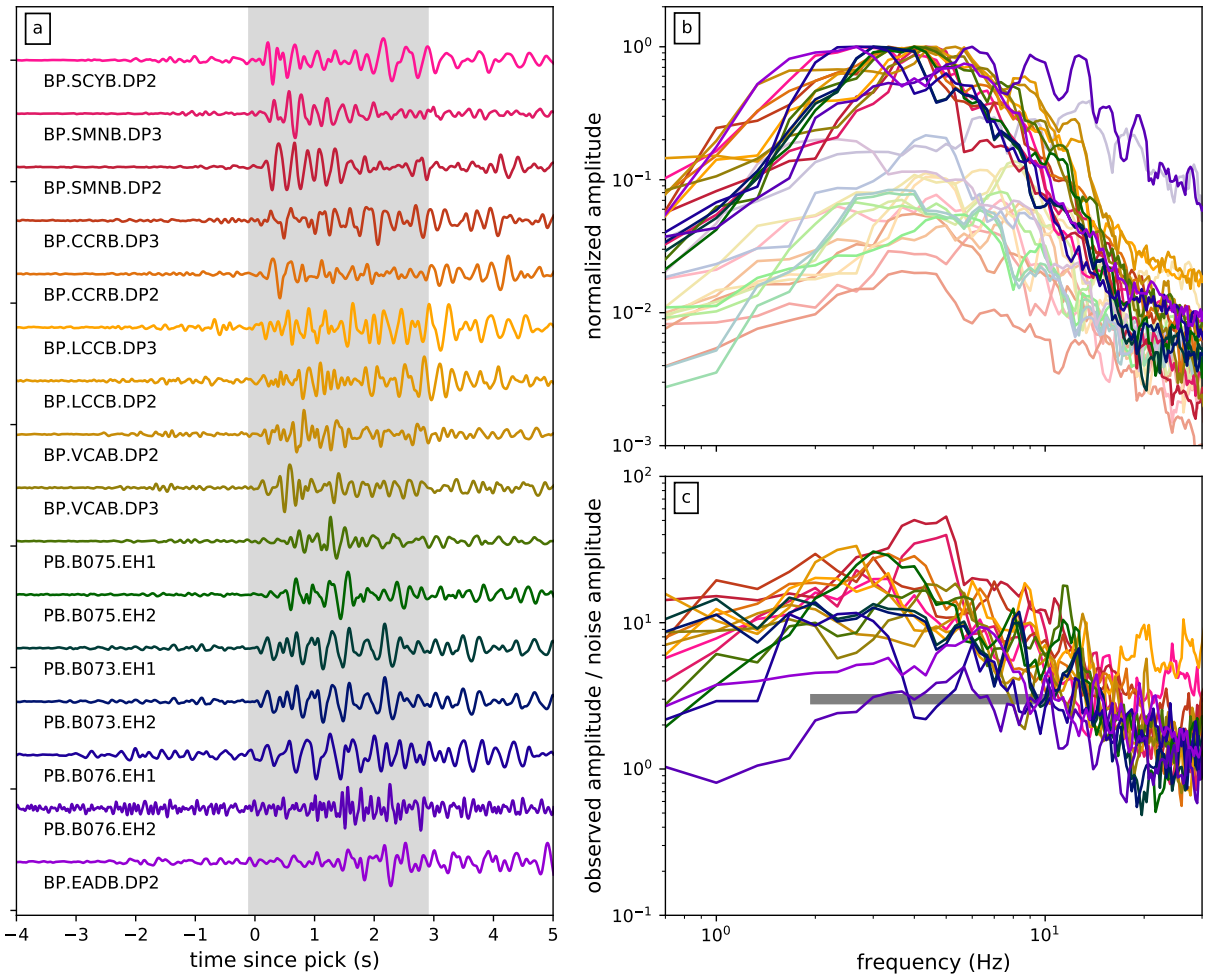

Figure S3: As in Figure S1, but for LFEs in family 9707.

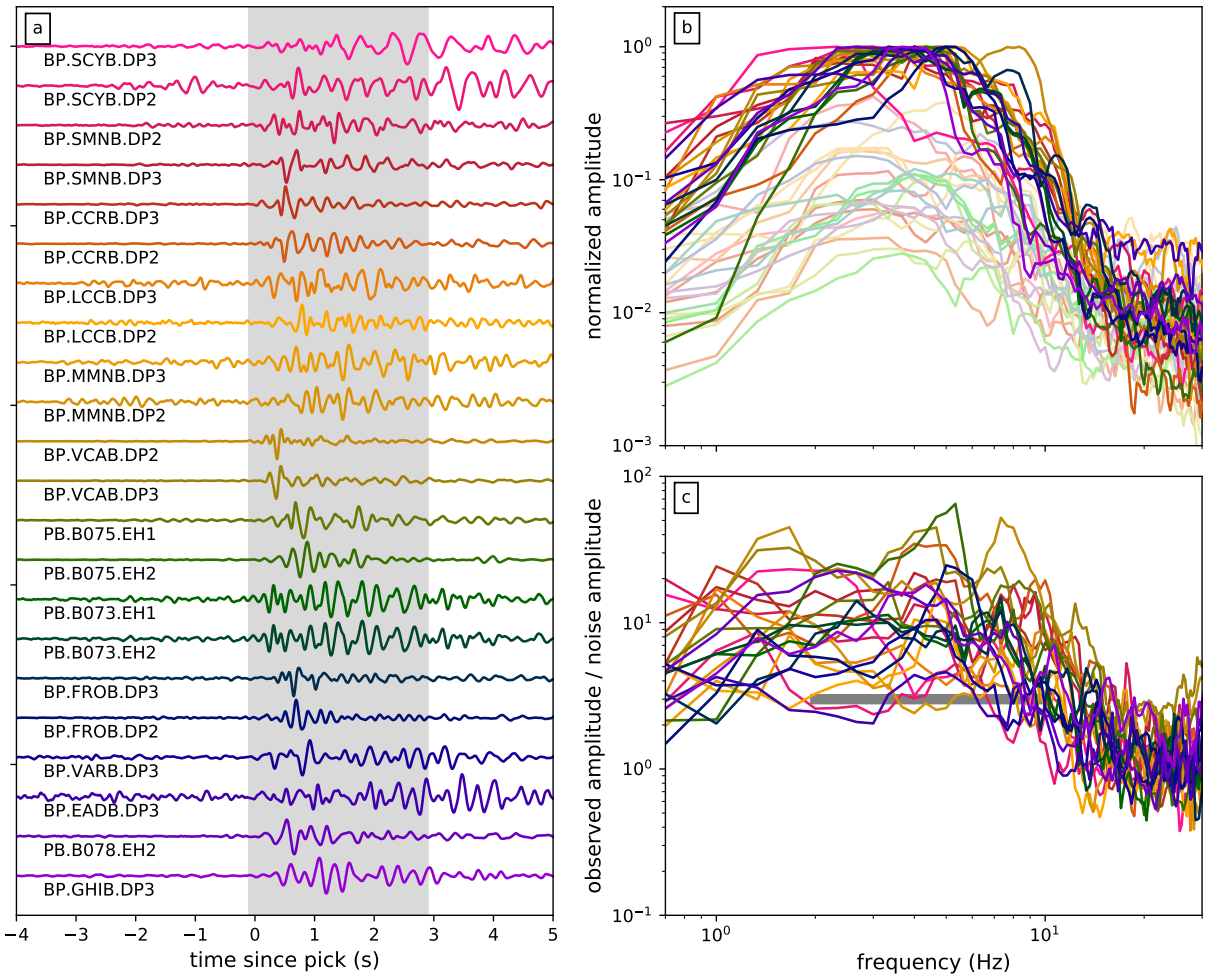

Figure S4: As in Figure S1, but for LFEs in family 77401.

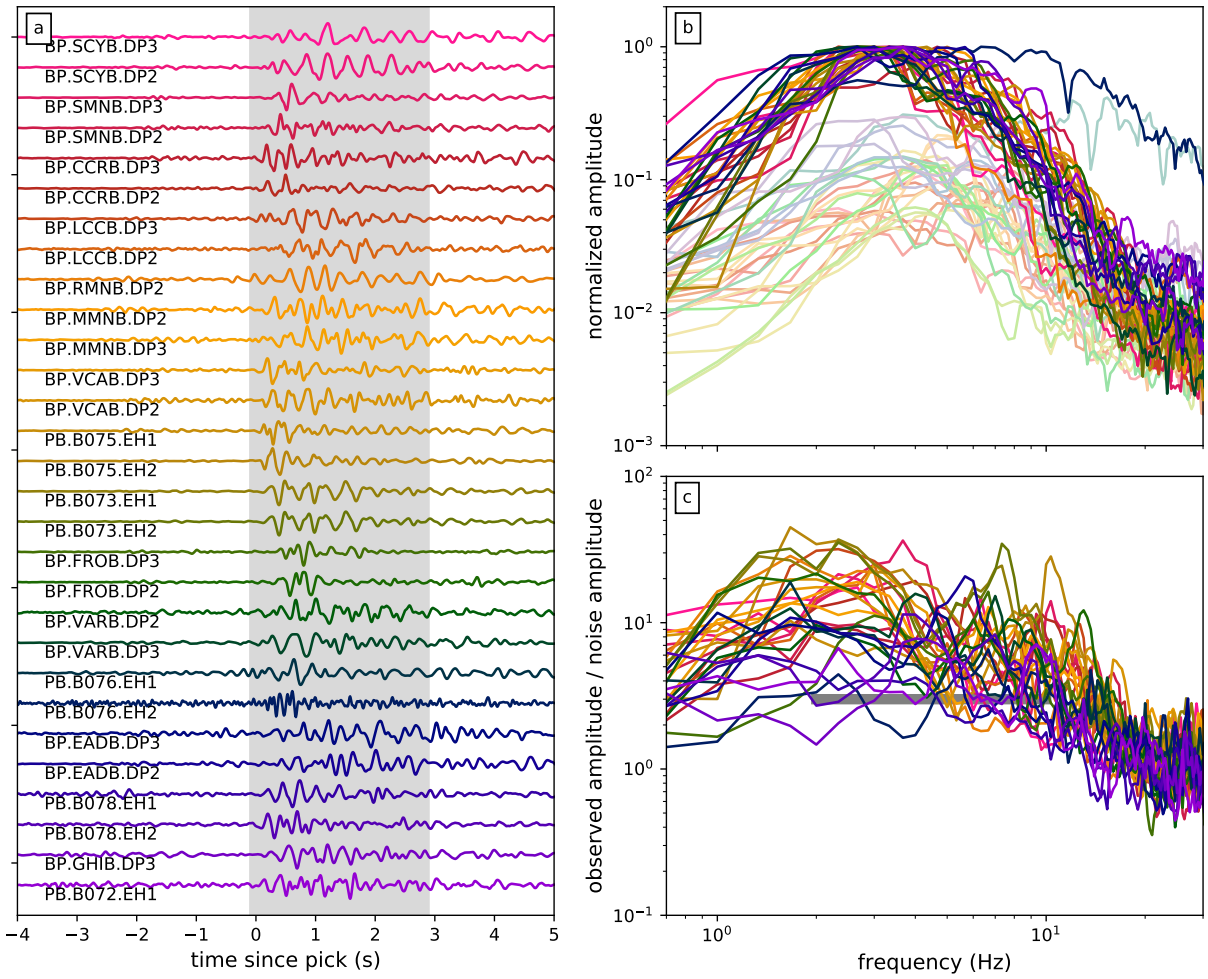

Figure S5: As in Figure S1, but for LFEs in family 27270.

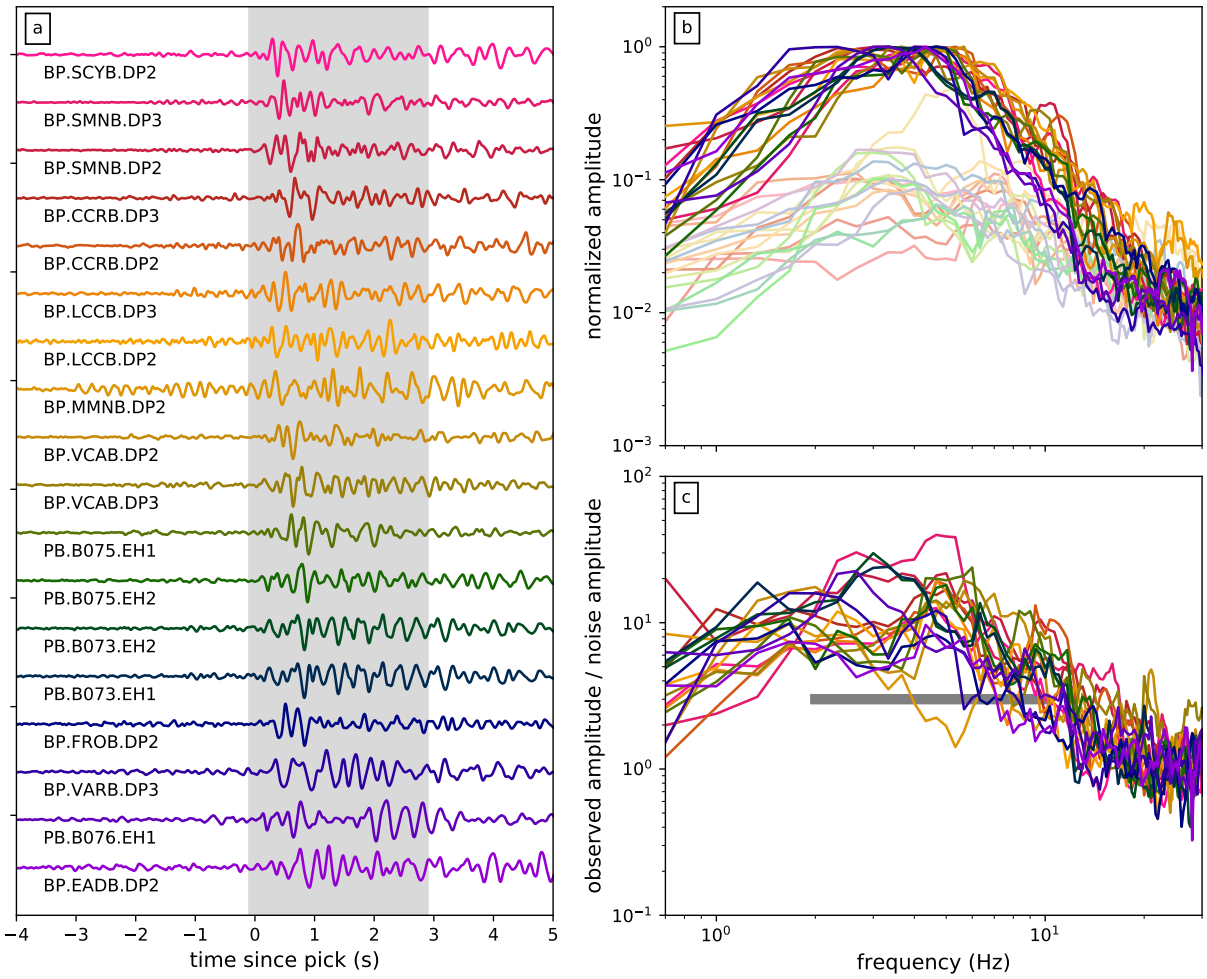

Figure S6: As in Figure S1, but for LFEs in family 45688.

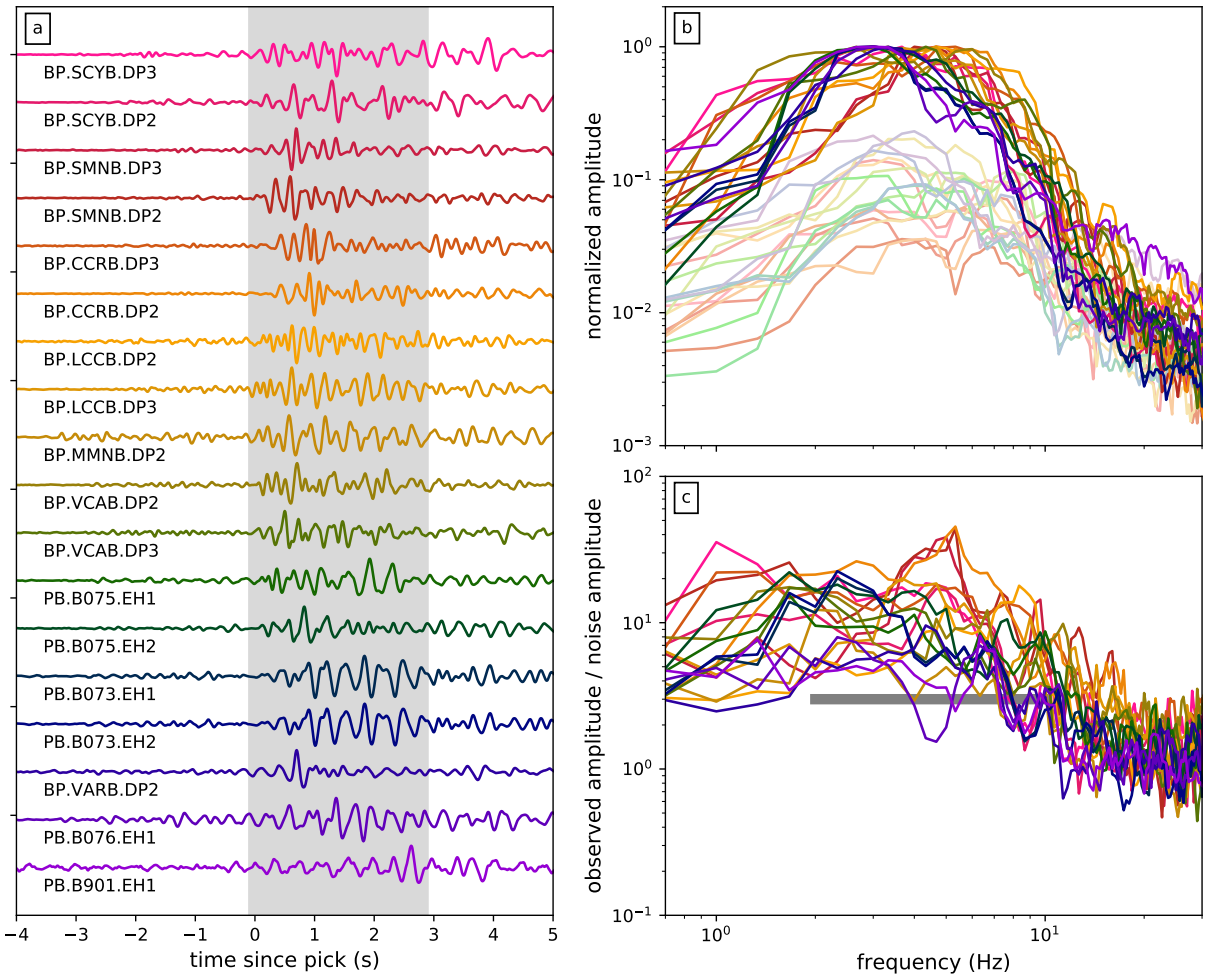

Figure S7: As in Figure S1, but for LFEs in family 70316.

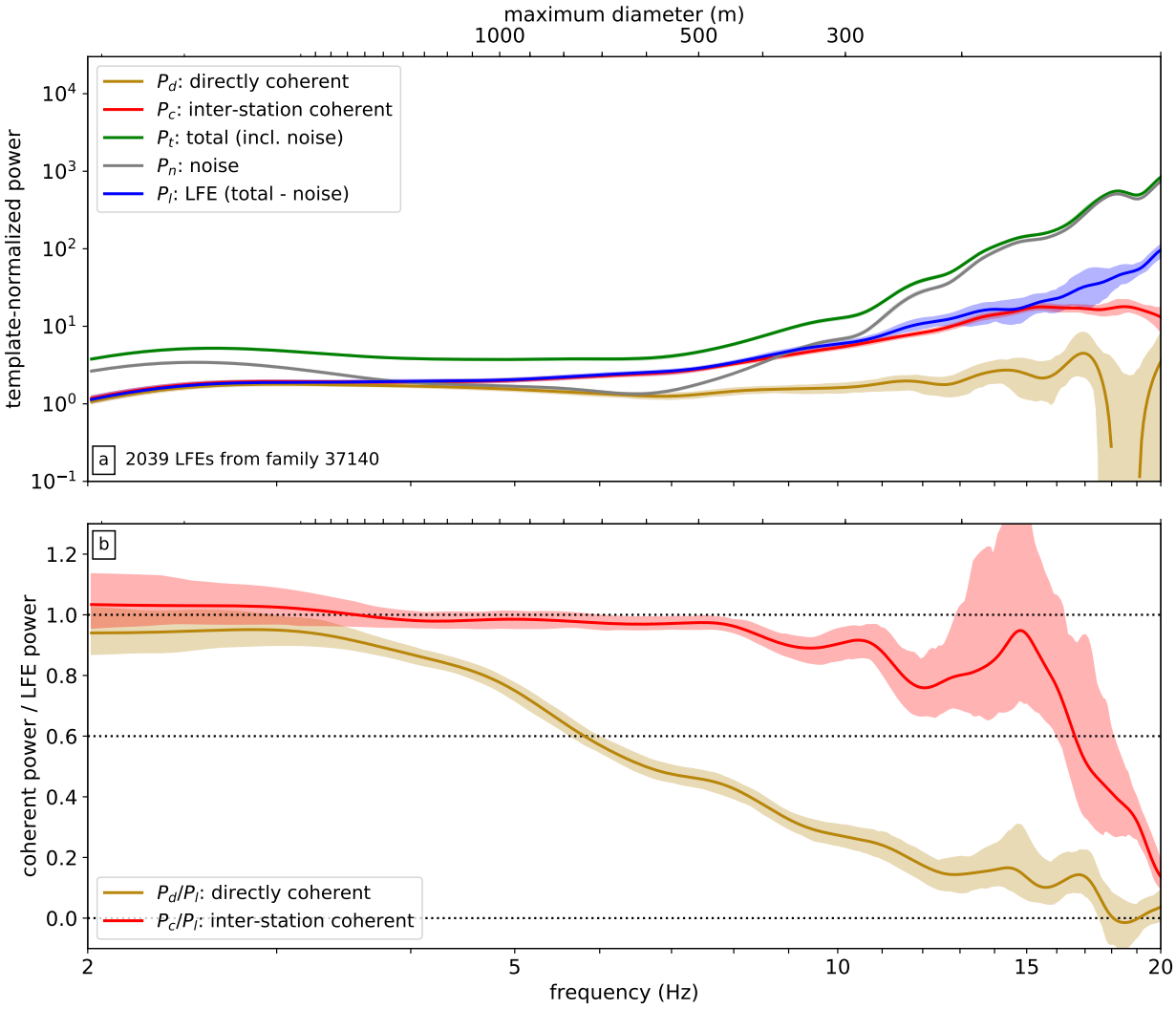

Figure S8: Coherent and total power for family 37140. Markings are as in Figure 4a and b.

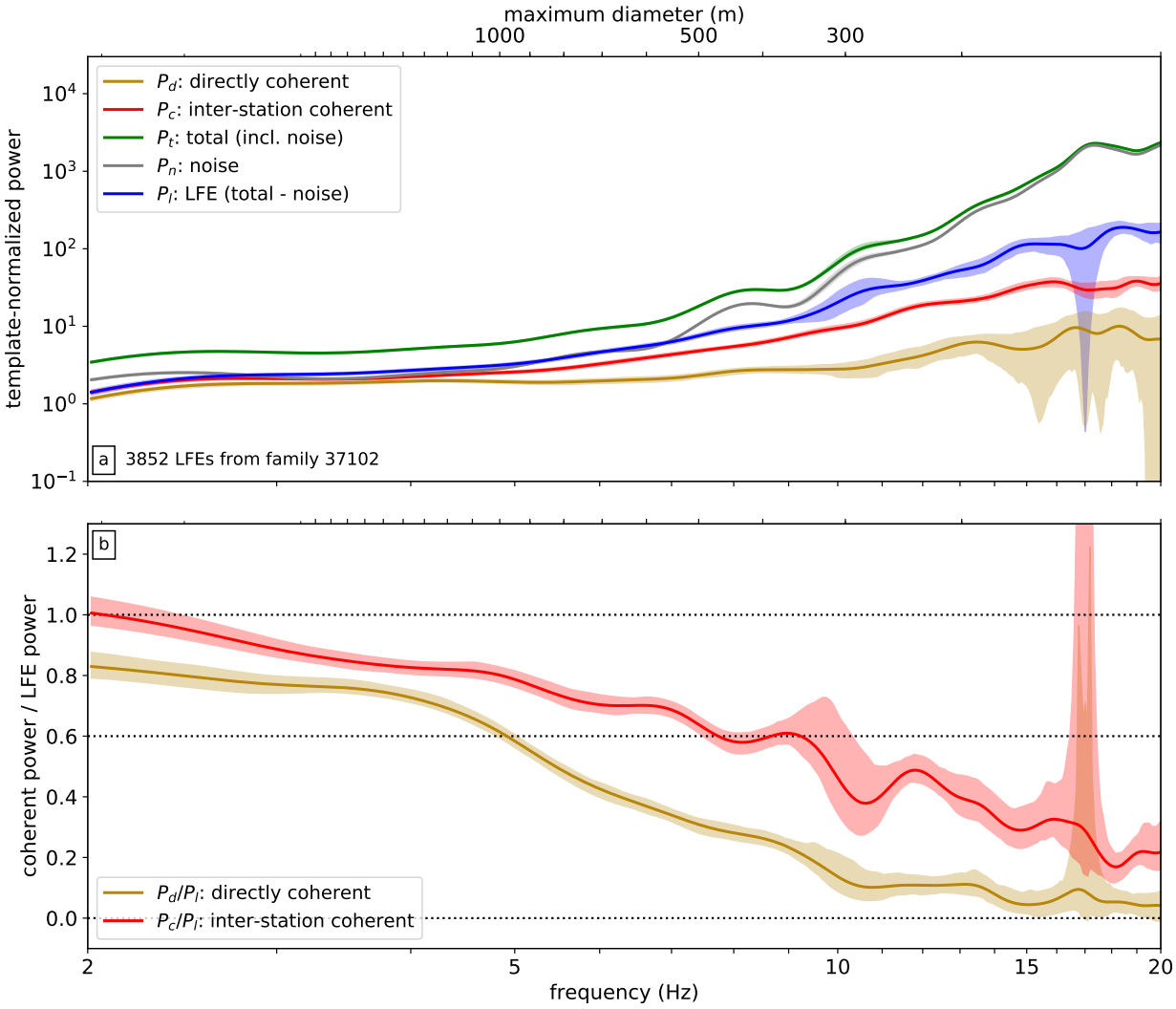

Figure S9: Coherent and total power for family 37102. Markings are as in Figure 4a and b.

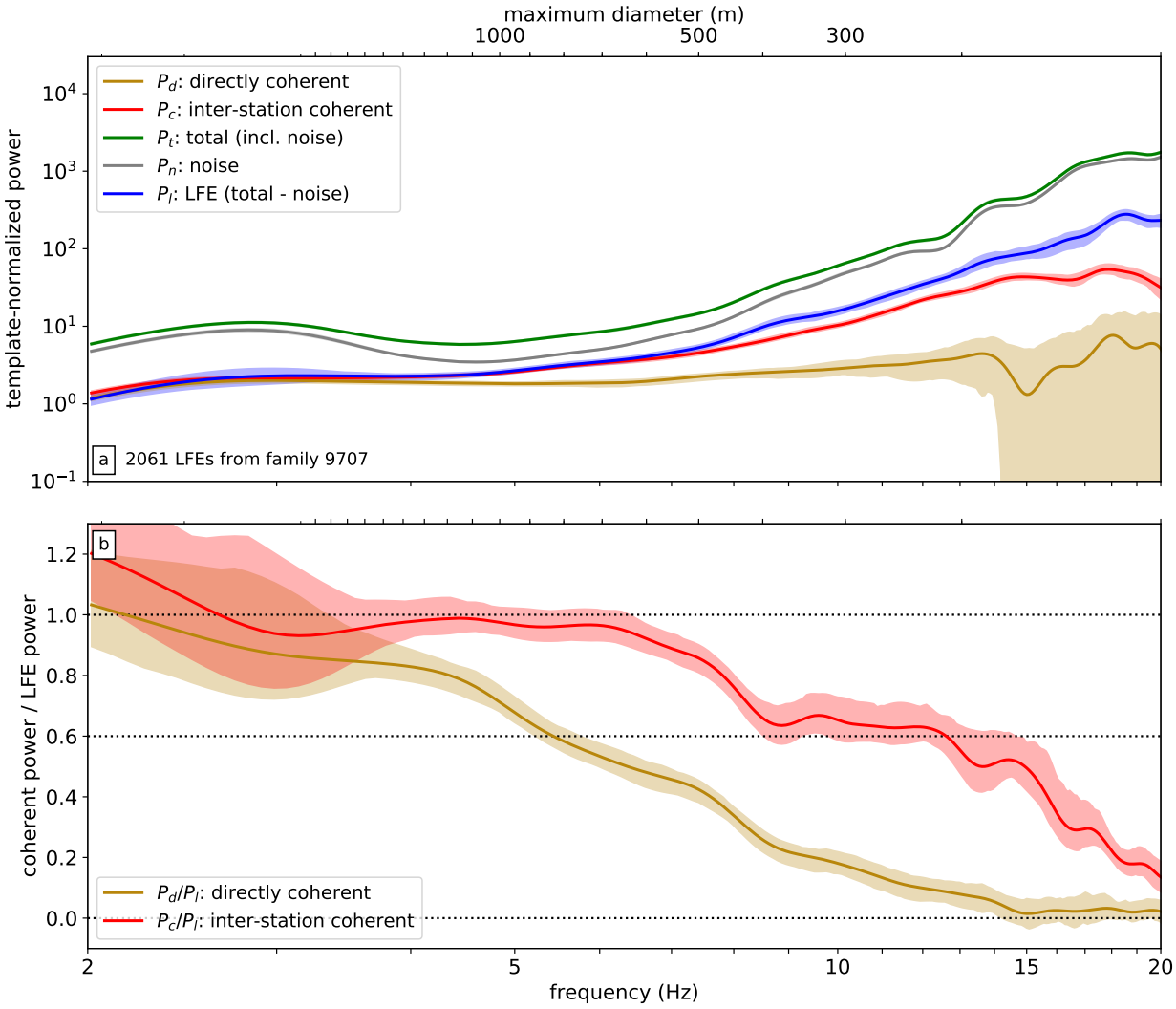

Figure S10: Coherent and total power for family 9707. Markings are as in Figure 4a and b.

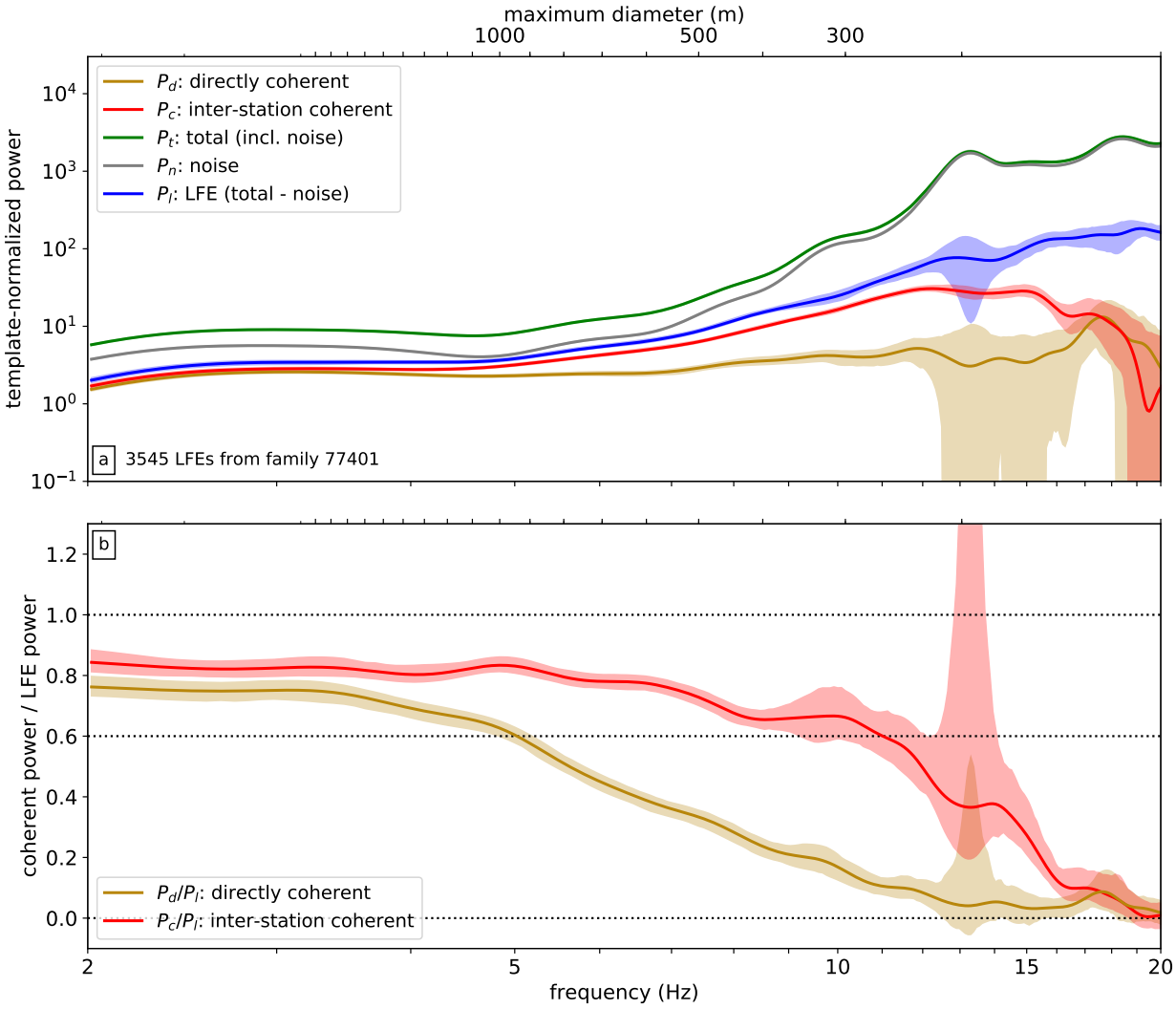

Figure S11: Coherent and total power for family 77401. Markings are as in Figure 4a and b.

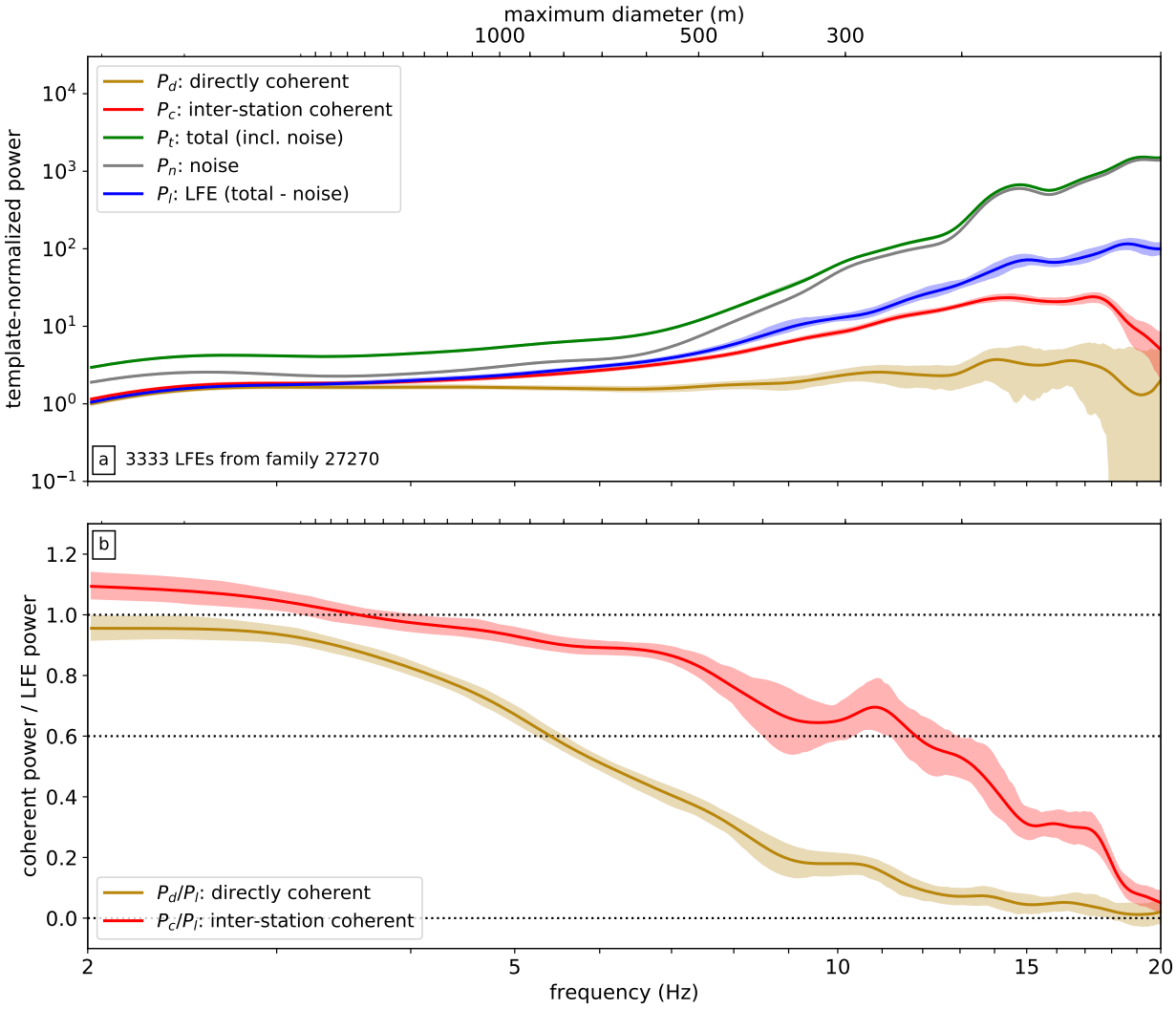

Figure S12: Coherent and total power for family 27270. Markings are as in Figure 4a and b.

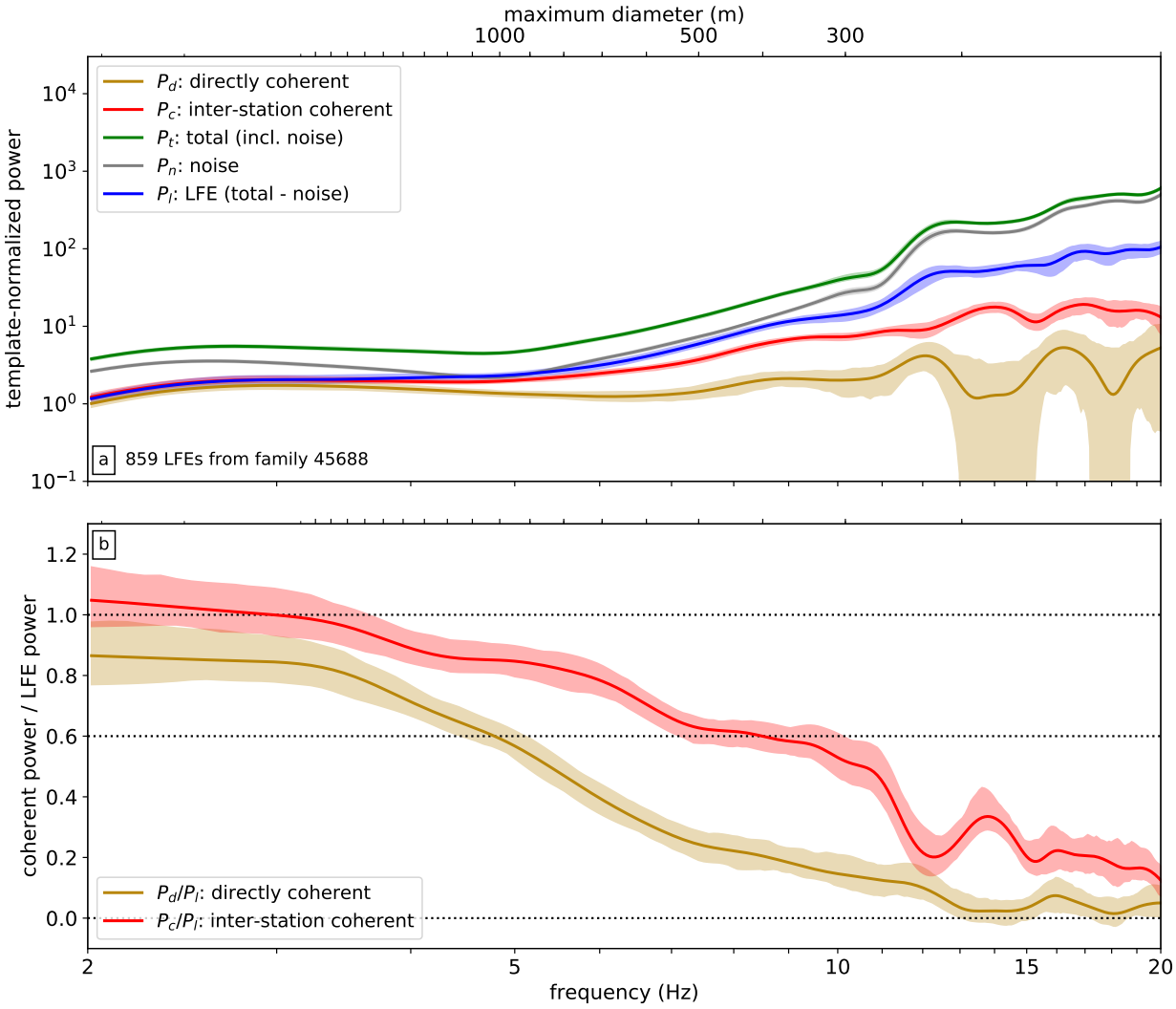

Figure S13: Coherent and total power for family 45688. Markings are as in Figure 4a and b.

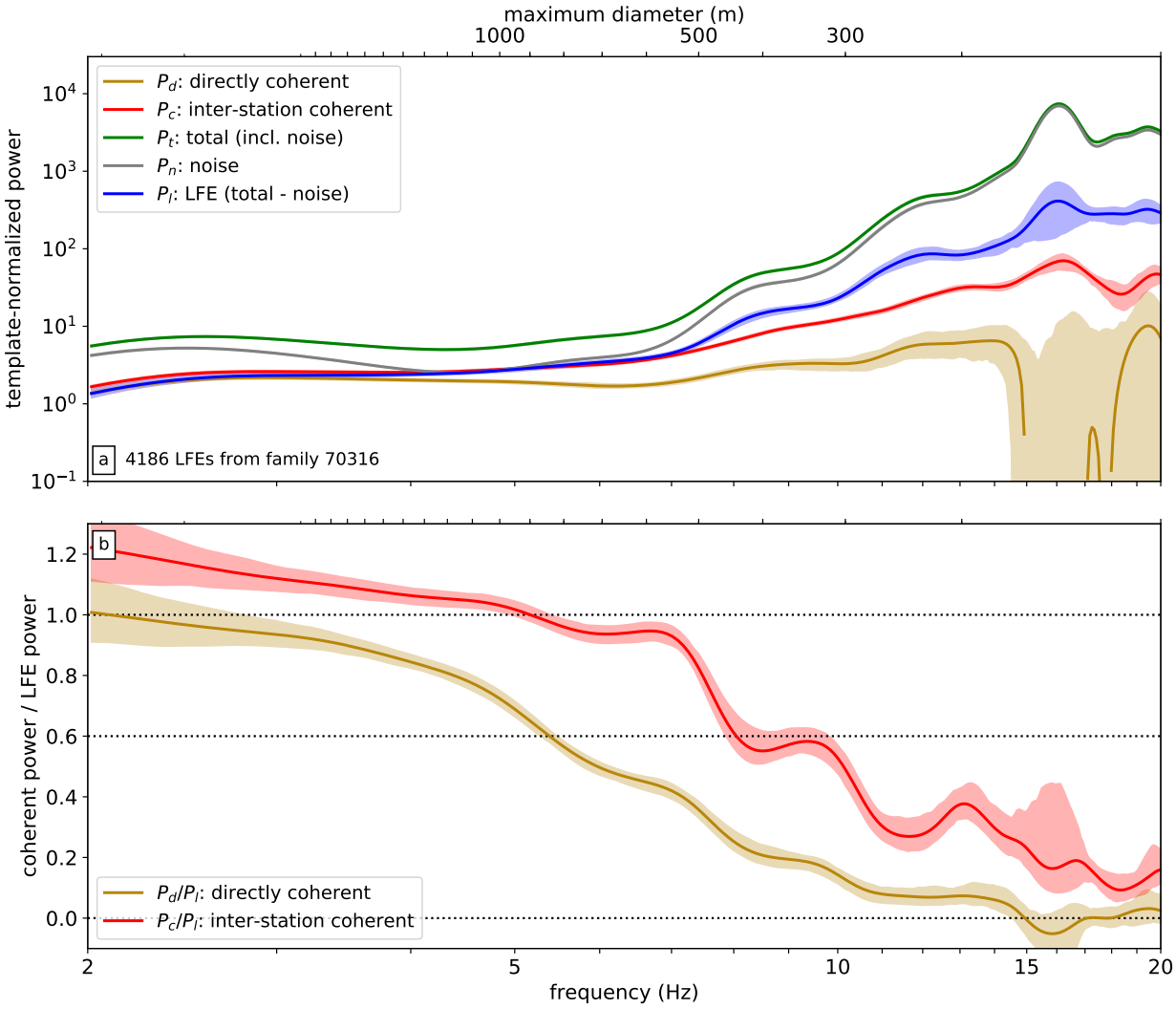

Figure S14: Coherent and total power for family 70316. Markings are as in Figure 4a and b.

S3 Coherence After Random Location Shifts

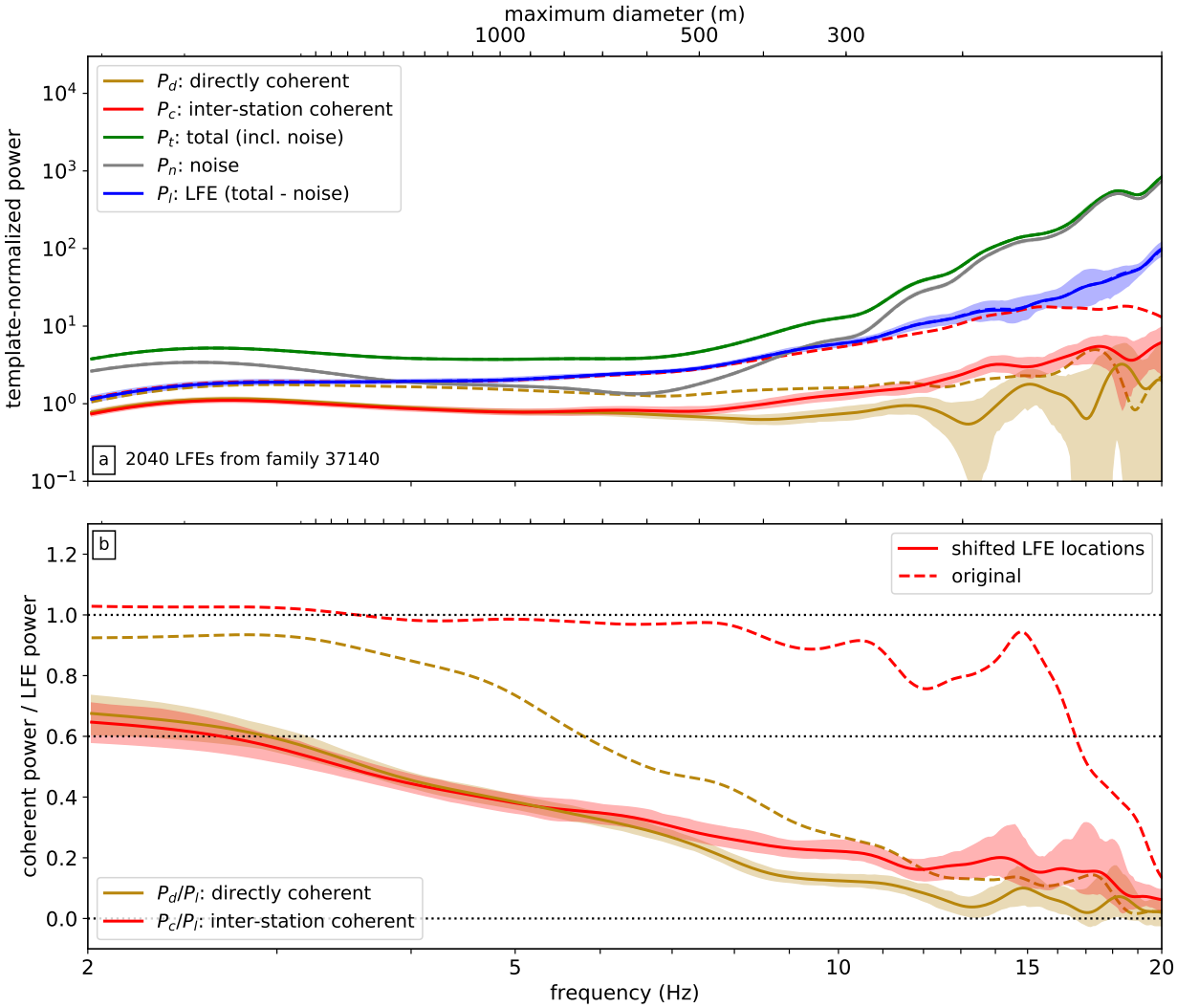

Figure S15: Coherent and total power for family 37140, for 1-km standard deviation shifts.

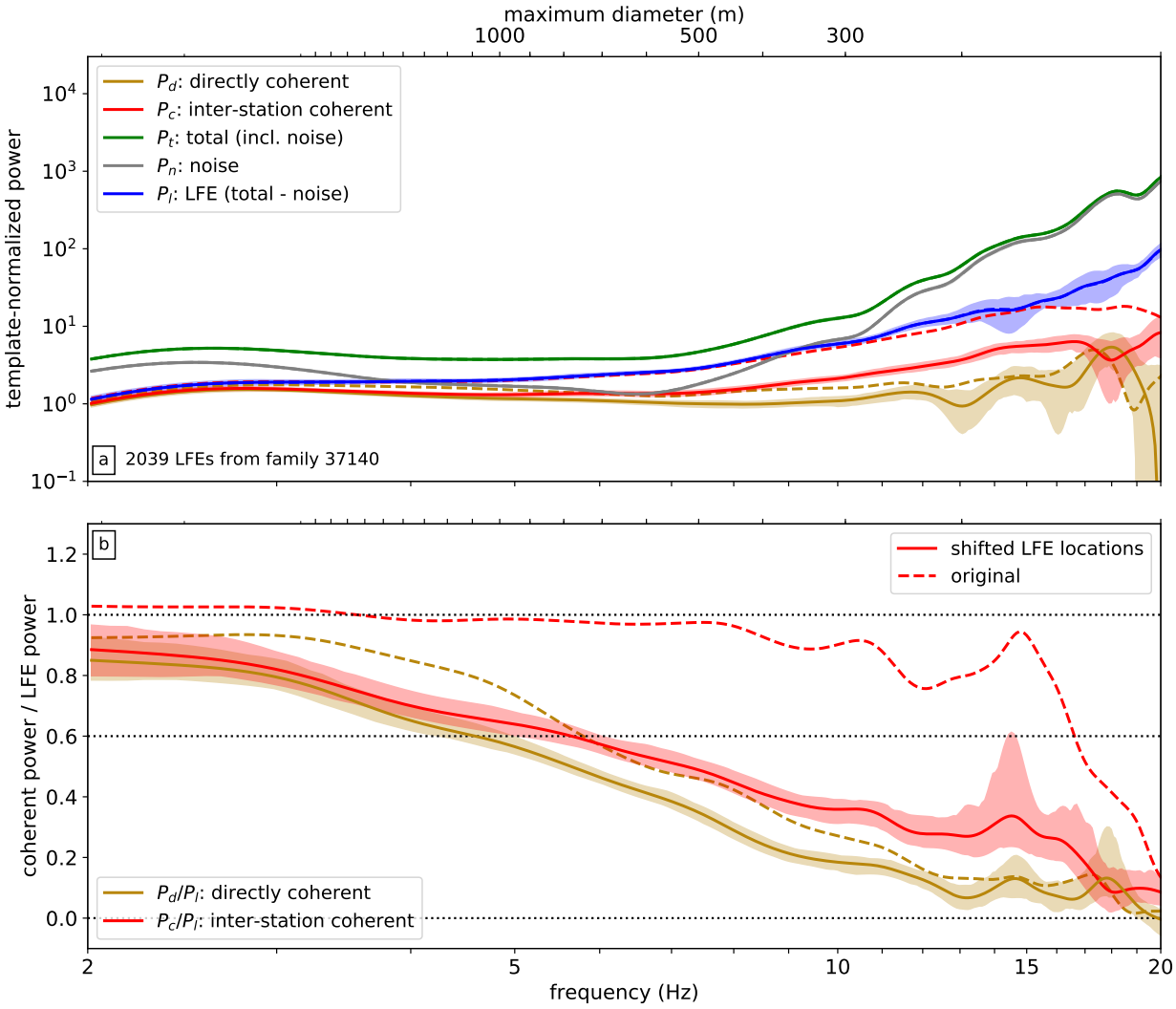

Figure S16: Coherent and total power for family 37140, for 0.75-km standard deviation shifts.

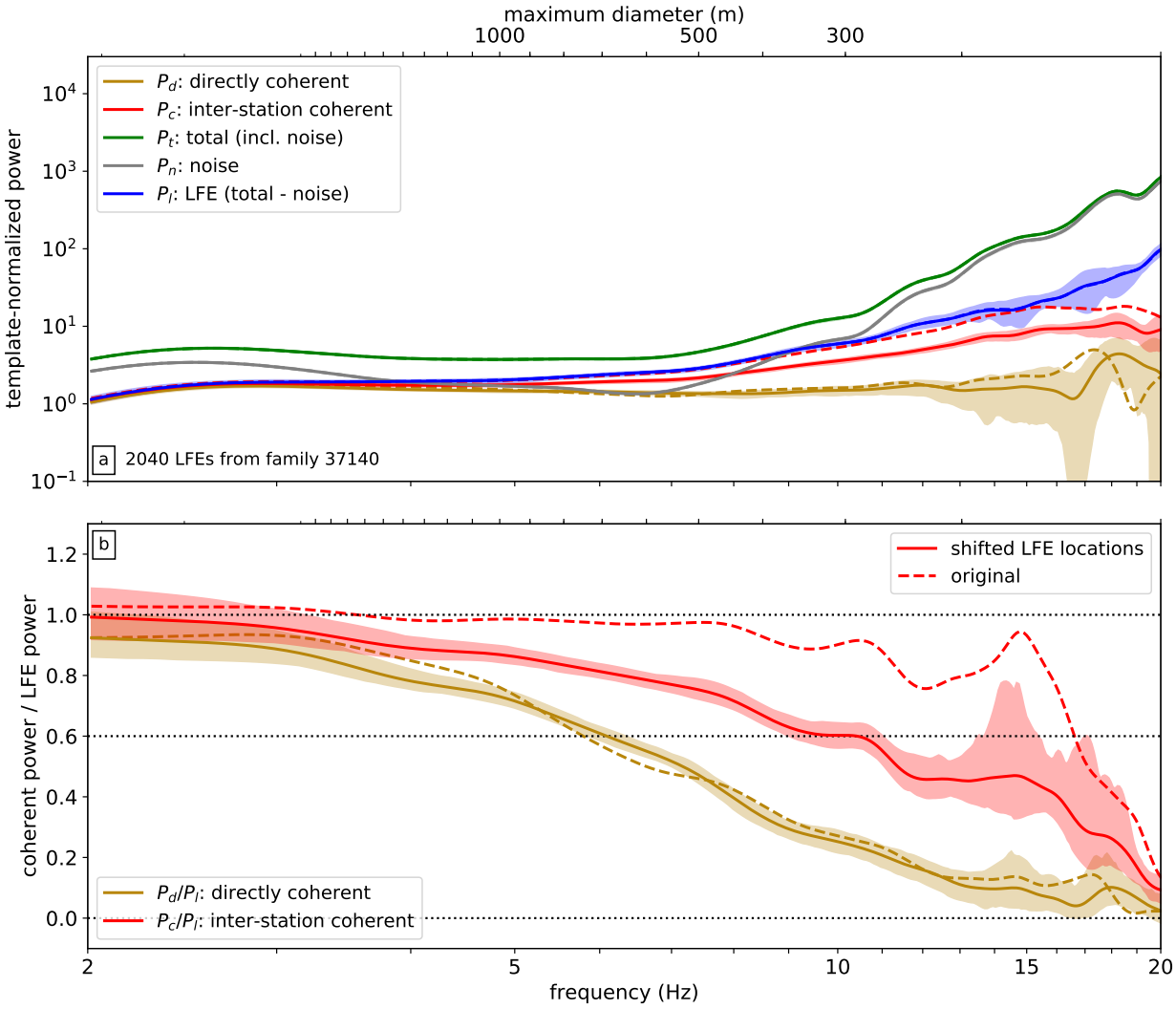

Figure S17: Coherent and total power for family 37140, for 0.25-km standard deviation shifts.

S4 Details of the Synthetic Rupture Calculations

In the synthetic LFE calculations described in section 6.1, we generate groups of 100 LFEs for various diameters D, rupture velocities V_r, rise times t_r, and complexities. Each LFE is given a roughly circular rupture area. The radius is allowed to vary randomly by up to 20% with azimuth. The diameter chosen from a lognormal distribution with factor of 1.3 standard deviation, centered at the specified average diameter.

We start by generating a fractal slip distributions for each LFE, with amplitude spectra that decay as wavenumber^{-2}, as suggested by observations of slip in large earthquakes (*Frankel*, 1991; *Herrero and Bernard*, 1994; *Mai and Beroza*, 2002). We taper the slip, multiplying by an elliptical profile $(1 - r^2/R^2)^{1/2}$, where r is the distance from the LFE center and R is the LFE radius at the relevant azimuth. Then we set any the slip values that are negative to zero so that the net slip is positive.

The figures below show synthetics created with variations on our approach. Adjustments are described in the captions.

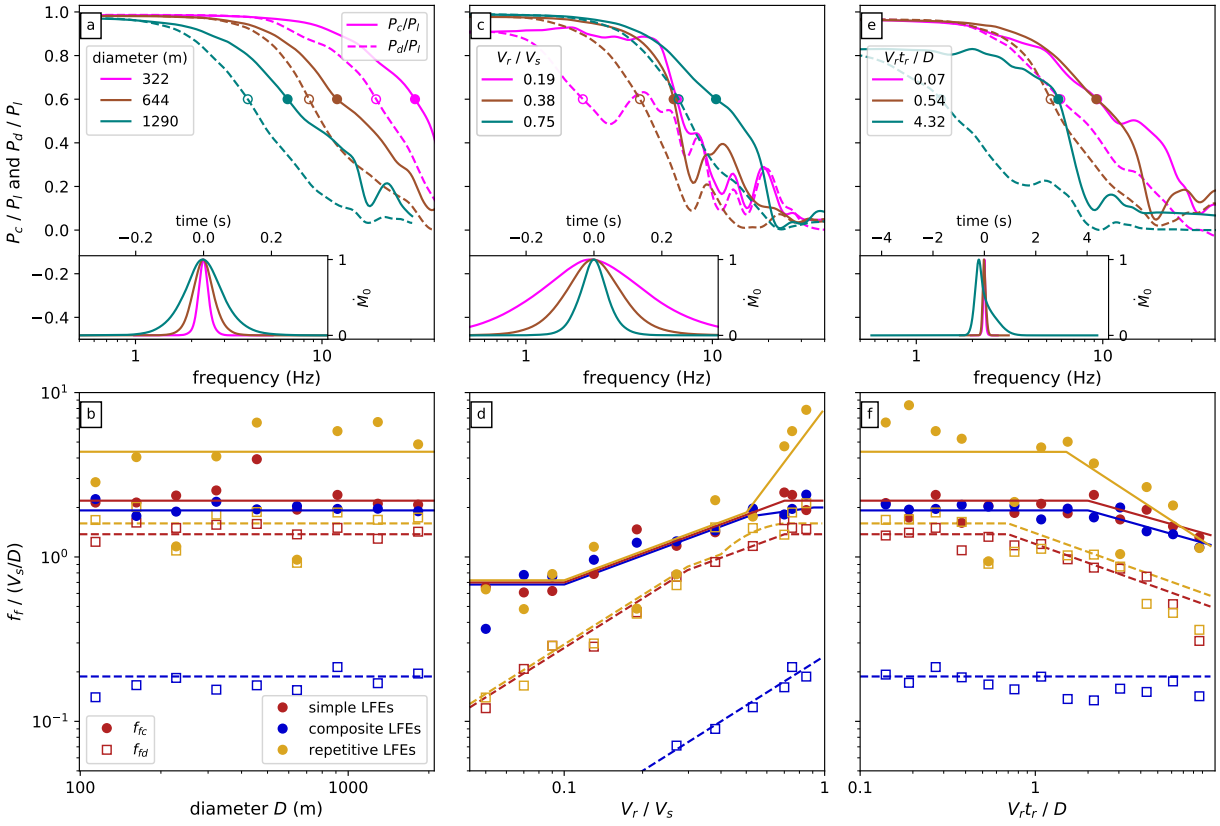

Figure S18: As in Figure 5, but using 500 LFEs in each group. Unlike in Figure 5, values in panels d and f are not averaged over a range of diameters.

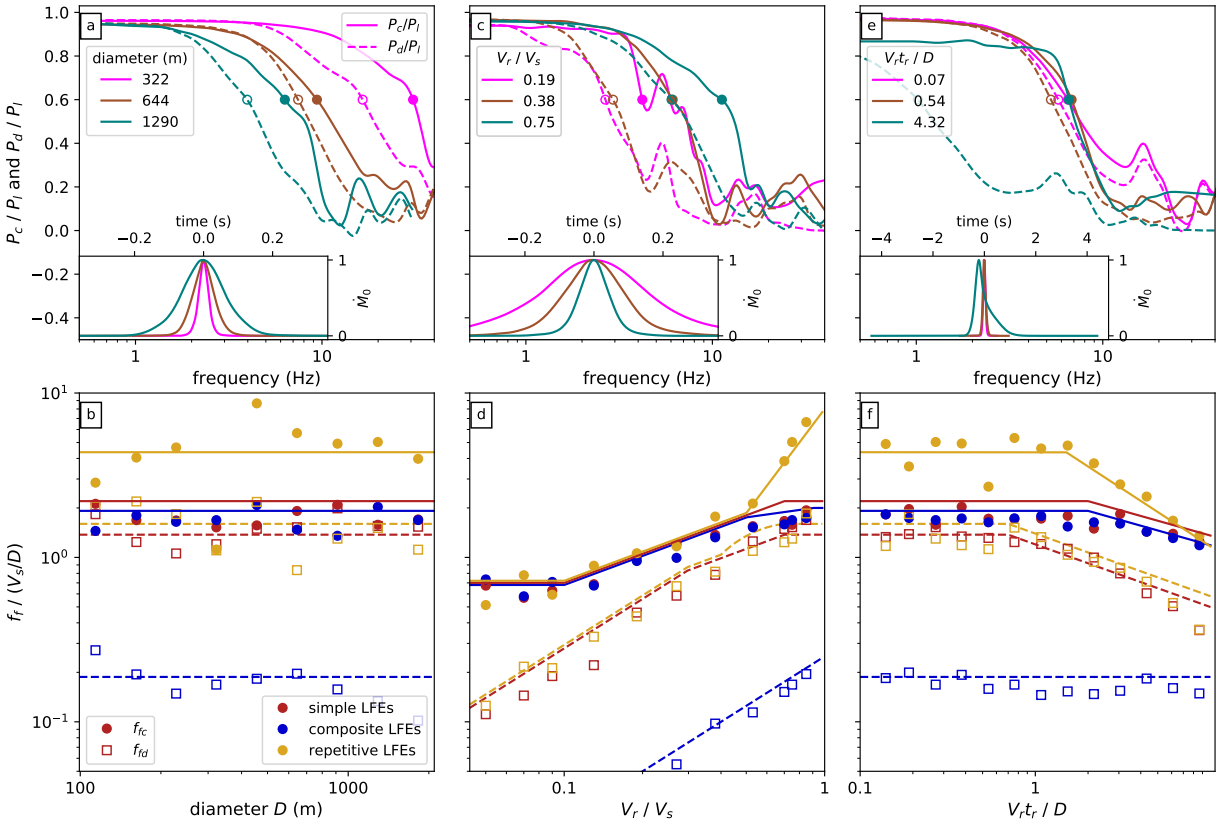

Figure S19: As in Figure 5, but here the simple LFEs (red circles and squares) are required to start within $0.1D$ of the rupture edge. The lines are as in Figure 5, chosen to match the LFEs that start at a random location within $0.4D$ of the rupture center, and thus plot slightly above the falloff frequencies obtained with the simulations shown.

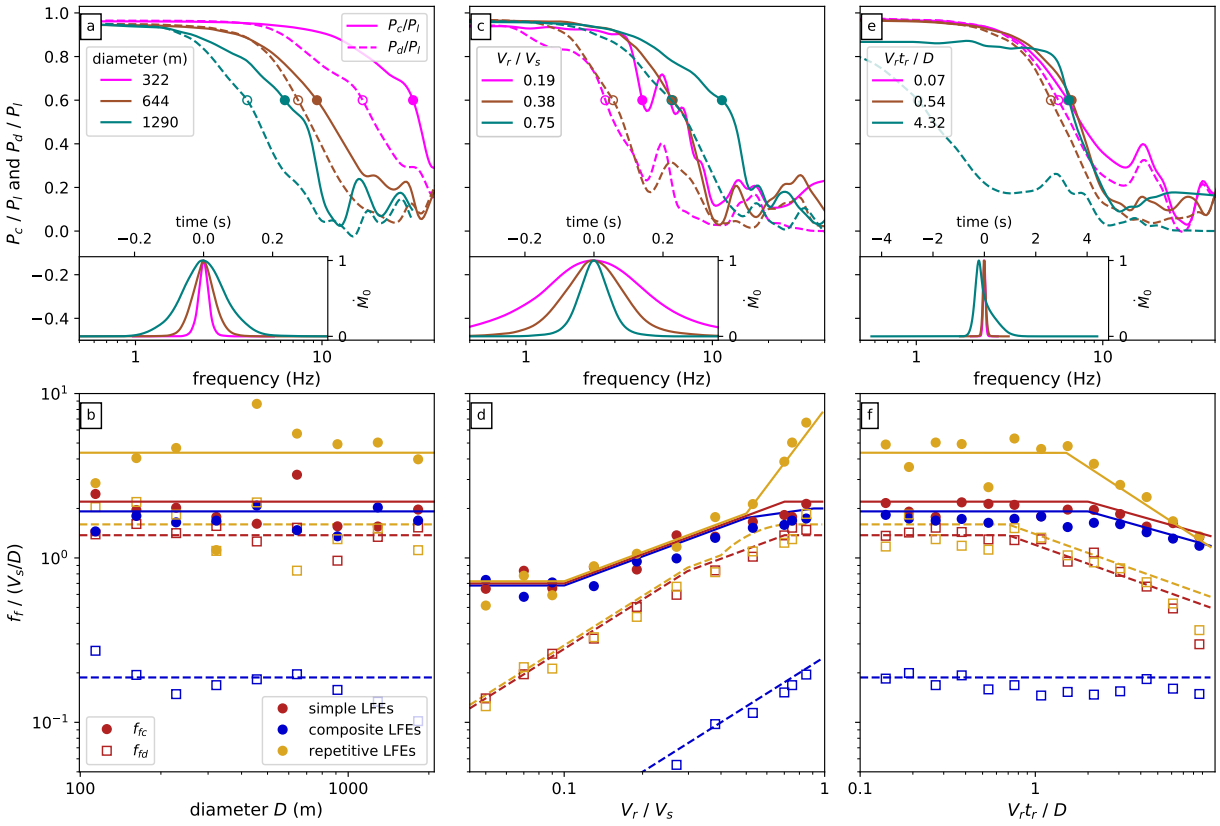

Figure S20: As in Figure 5, but here the simple LFEs (red circles and squares) are computed with a random subset of 5 stations, not with all the stations available for family 37140.

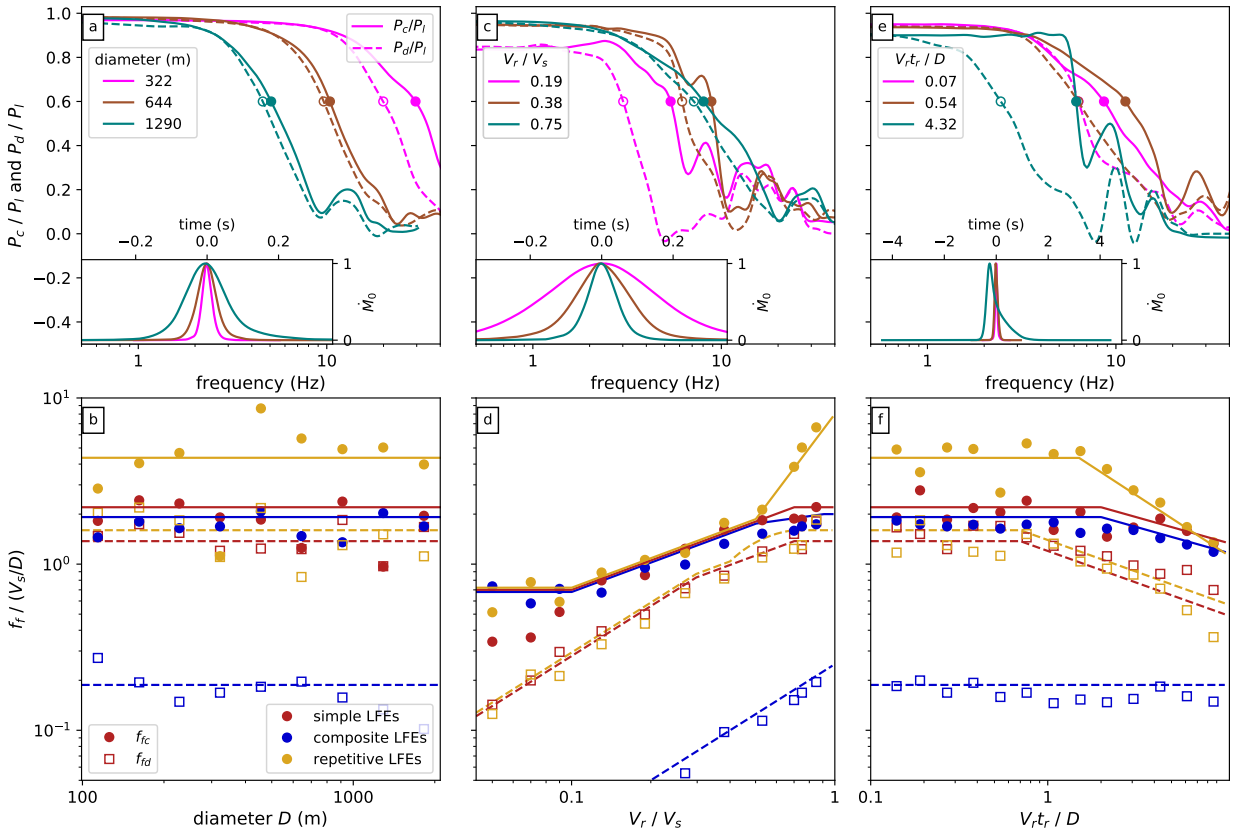

Figure S21: As in Figure 5, but here when the simple LFEs (red circles and squares) are computed, we assign the observing stations to random azimuths between 0 and 360°. We retain the takeoff angles for the stations recording family 37140 LFEs, however.

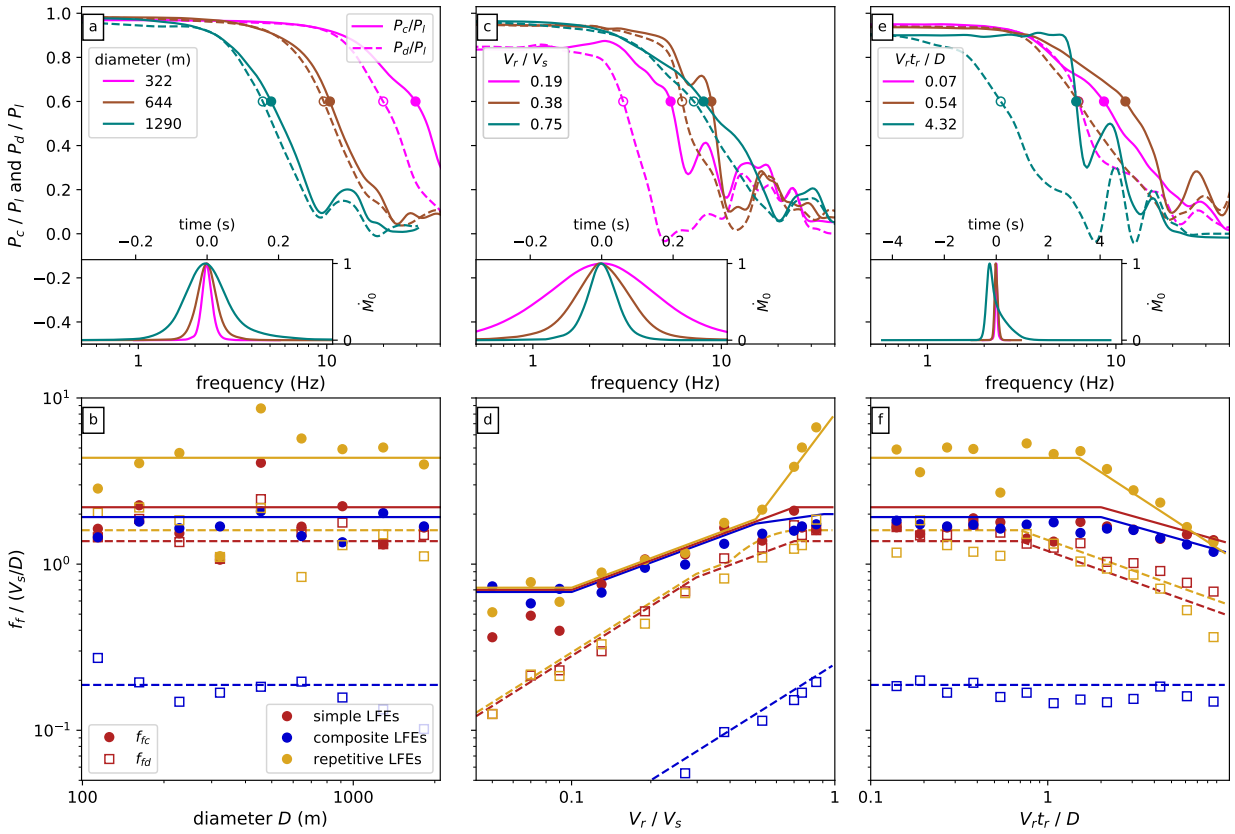

Figure S22: As in Figure 5, but here the simple LFEs (red circles and squares) are required to start within $0.1D$ of the rupture edge. In addition, the observing station locations are assigned random azimuths between 0 and 360°, as in Figure S21.

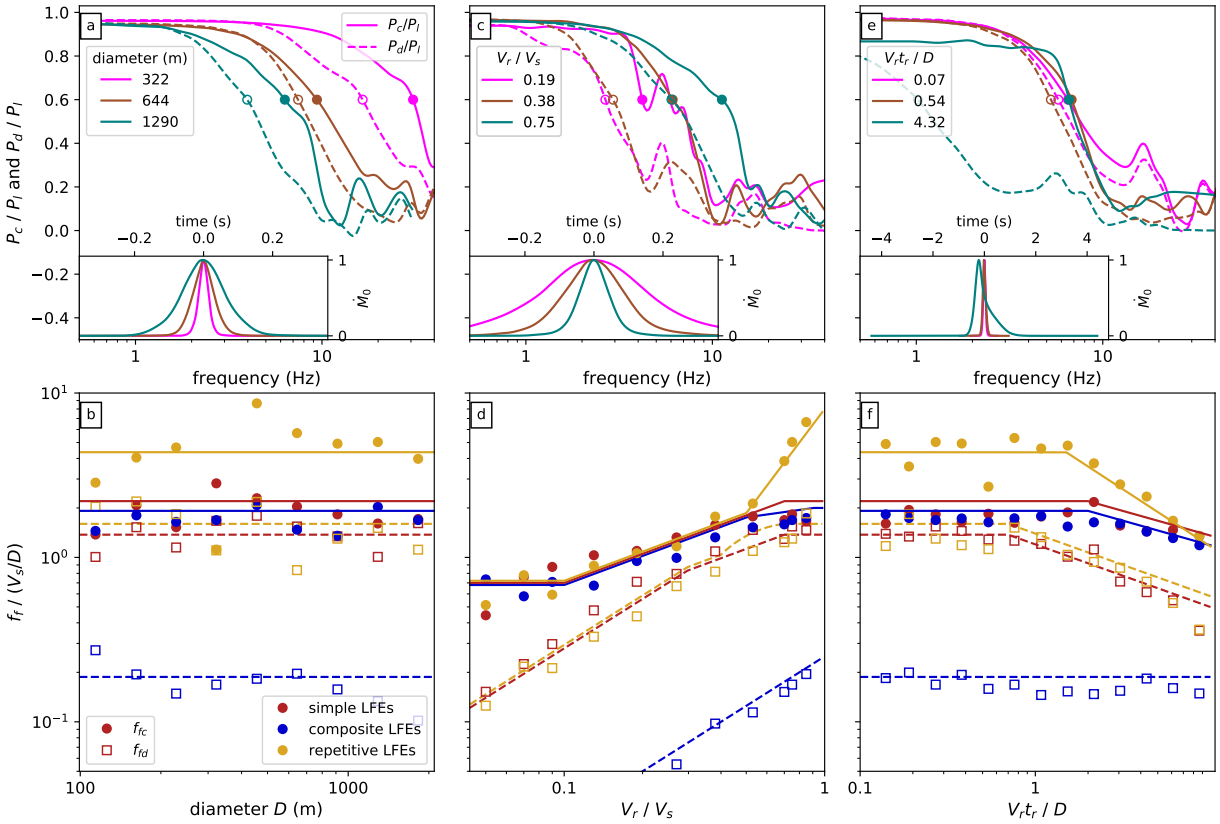

Figure S23: As in Figure 5, but here the simple LFEs (red circles and squares) are computed with slip distributions that are offset from zero before computation. We create fractal slip distributions and then shift the mean of the slip distribution so that 90% of the values are positive within the rupture area. Then we taper the slip and compute apparent source time functions.

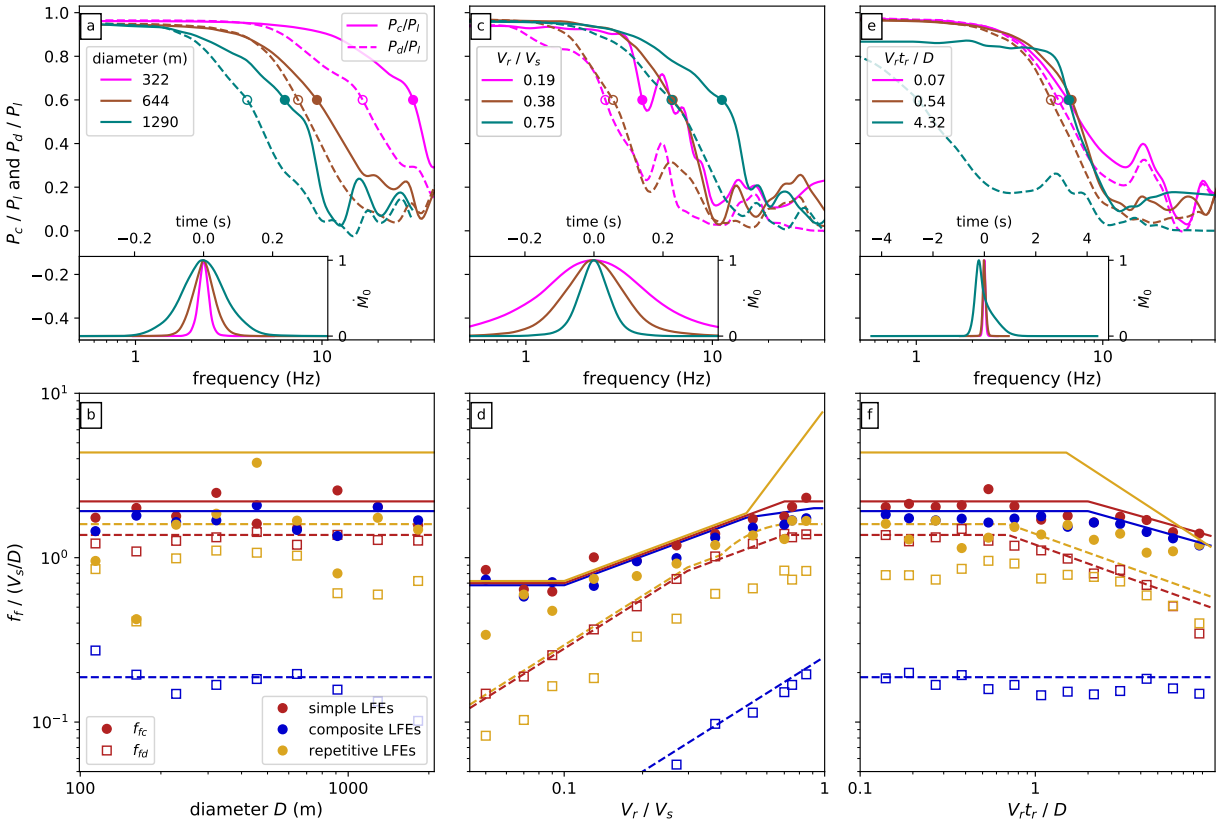

Figure S24: As in Figure 5, but here the repeating LFEs (yellow circles and squares) are computed with different directivity. About 75% of events in each group start on the SE edge, and 25% start on the NW edge.

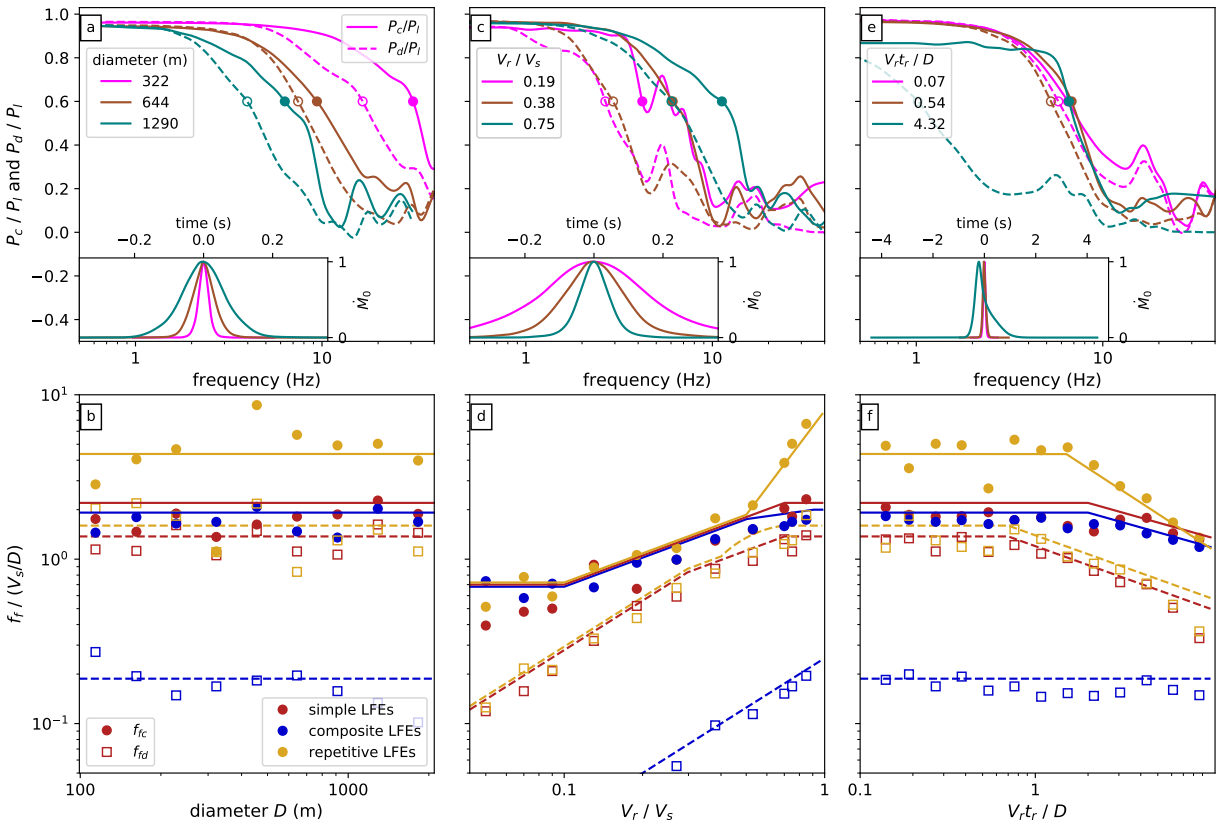

Figure S25: As in Figure 5, but the radii are chosen from lognormal distributions with factor of 1.1 standard deviation instead of factor of 1.3 standard deviation.

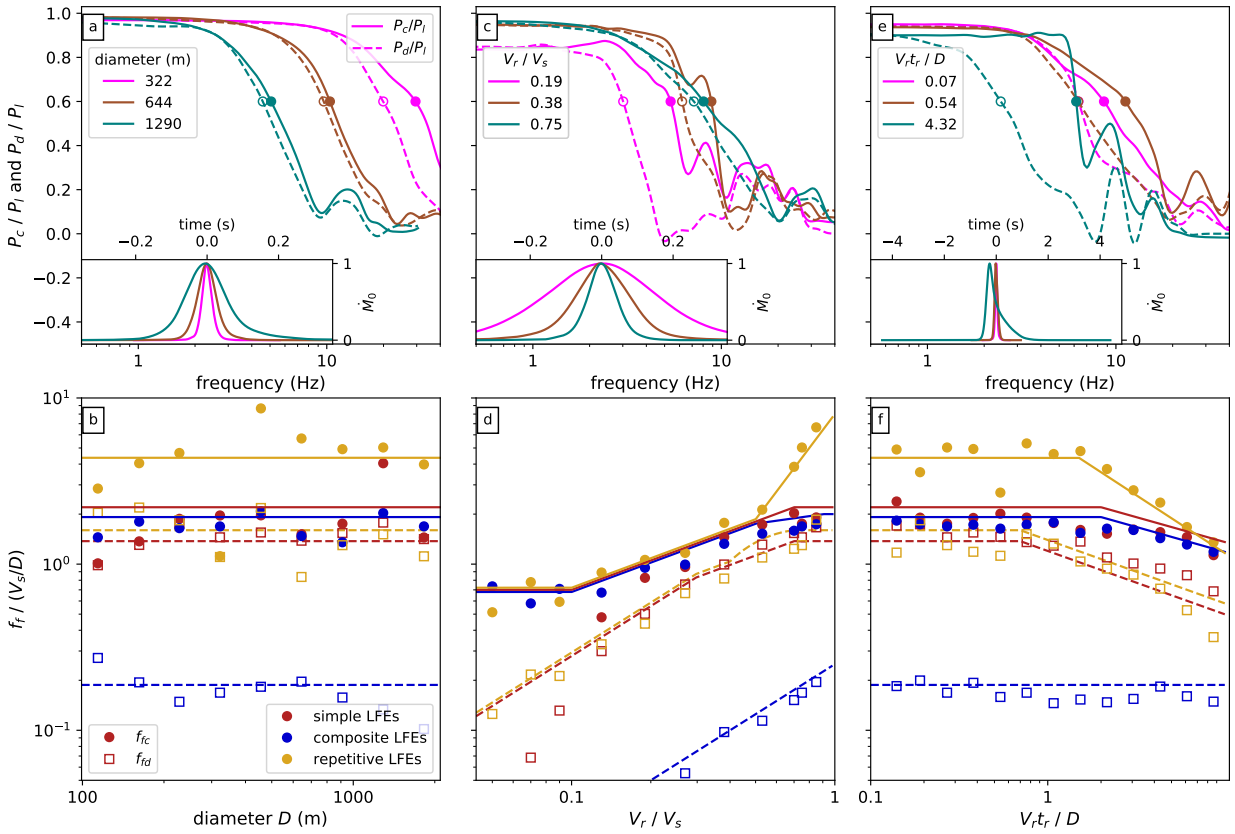

Figure S26: As in Figure 5, but here the Green's functions are not tapered white noise, but are instead taken to be the seismograms of a local earthquake that occurred 5 km away from the LFEs in family 37140.

References

Frankel, A. (1991), High-frequency spectral falloff of earthquakes, fractal dimension of complex rupture, b value, and the scaling of strength on faults, *J. Geophys. Res.*, *96*, 6291–6302, doi: 10.1029/91JB00237.

Hawthorne, J. C., and J.-P. Ampuero (2017), A phase coherence approach to identifying co-located earthquakes and tremor, *Geophys. J. Intern.*, *209*(2), 623–642, doi:10.1093/gji/ggx012.

Herrero, A., and P. Bernard (1994), A kinematic self-similar rupture process for earthquakes, *Bull. Seis. Soc. Amer.*, *84*(4), 1216–1228.

Mai, P. M., and G. C. Beroza (2002), A spatial random field model to characterize complexity in earthquake slip, *J. Geophys. Res.*, *107*(B11), 2308, doi:10.1029/2001JB000588.

Shelly, D. R. (2017), A 15 year catalog of more than 1 million low-frequency earthquakes: Tracking tremor and slip along the deep San Andreas Fault, *J. Geophys. Res.*, *122*(5), 3739–3753, doi:10.1002/2017JB014047.